CS 134: Networks	Section 6
Prof. Yaron Singer	Spring 2017

1 Submodular Games

Recall from lecture the approximation algorithm we gave.

Algorithm 1 Greedy Algorithm

```
1: Set S = ∅
2: while |S| ≤ k do
3:     S ← S ∪ argmax_{a∈N} f_S(a)
4: end while
5: return S
```

In the proof for this algorithm, we made use of 2 properties of submodular functions:

- If f, g are monotone submodular functions, and $\alpha, \beta > 0$ then:

$$h(S) = \alpha f(S) + \beta g(S)$$

  is a monotone submodular function as well

- a function $f : 2^N \to \mathbb{R}$ is submodular if and only if for every $S \subseteq N$ the marginal contribution function $f_S(T) = f(S \cup T) - f(T)$ is subadditive

Some useful definitions:

Definition. *A function $f : 2^N \to \mathbb{R}$, is* ***monotone*** *if $S \subseteq T \implies f(S) \leq f(T)$.*

Definition. *Given a function $f : 2^N \to \mathbb{R}$, the* ***marginal contribution*** *of an element $e \in N$ to $S \subseteq N$ is $f_S(e) = f(S \cup e) - f(S)$.*

Definition. *A function $f : 2^N \to \mathbb{R}$ is* ***submodular*** *if for any $e \in N$ we have that:*

$$f_S(e) \geq f_T(e) \quad \forall S \subseteq T \subseteq N$$

Definition. *A function $f : 2^N \to \mathbb{R}$ is* ***subadditive*** *if for any $S, T \subseteq N$ we have that:*

$$f(S \cup T) \leq f(S) + f(T)$$

Exercise

Prove these 2 properties

2 Random Walks

adapted from https://people.csail.mit.edu/ronitt/COURSE/S08/download/notes15.pdf

In lecture we examined voter influence through transition matrices. By taking powers of these matrices, we simulated "time steps" through which influence spread. Closely related to that model is the

Random Walk: given an undirected graph and a starting node, what sorts of paths do we see after t time steps?

We will be looking at the expected time to reach a node from a given starting node. We define this to be the hitting time from node i to j, denoted h_{ij}. We define the cover of a connected, undirected graph G to be

$$\mathcal{C}(G) = \max_u C_u(G)$$

where,

$$\mathcal{C}_u(G) = E[\#\text{ of steps to reach all nodes in } G \text{ on a walk that starts at } u]$$

We will prove that $O(n^3)$ is the worst possible cover time. We first add in the assumption that G is aperiodic. This means that there is no integer $k > 1$ that divides the length of every cycle in the graph. This can be done by adding or removing self-loops.

We define the commute time from i to j as C_{ij}: the expected number of steps for a random walk starting at i to reach j and then return to i. So, $C_{ij} = h_{ij} + h_{ji}$

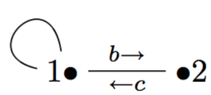
Lemma 1. *$\forall (u, v) \in E$, we have $C_{uv} \leq 2m$*

Proof. Consider a walk of the form $v \to u \rightsquigarrow v \to u$. We'll show that

$$\mathbb{E}[\text{time between 2 visits to the directed edge } (v, u)] \leq 2m$$

This bounds C_{uv}. If we're at u and we assume that we just came from v, then after we visit $v \to u$ again, we've gone from u to v and back to u.

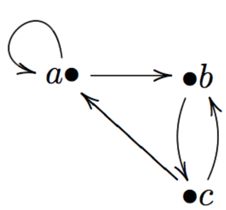
To do this, construct $G' = (V', E')$ from $G = (V, E)$, which represents walking on edges of G. V' is the set of directed edges in G, that is, for every undirected edge between x and y in E, we have two edges (x, y) and (y, x) in V'. The set of edges is $E' = \{((u, v), (v, w)) | (u, v), (v, w) \in V'\} \subseteq V'^2$. For example, the graph, transition matrix, and walk



A	1	2
1	1/2	1/2
2	1	0

$1 \to 1 \to 2 \to 1 \to 1$

become



A	a	b	c
a	1/2	1/2	0
b	0	0	1
c	1/2	1/2	0

$a \to b \to c \to a$

G' is called the "line graph" of G. Our transformed goal is now to find the hitting time of $h_{(u,v),(u,v)}$.

Observe that G' is doubly stochastic since $P'_{(u,v)(v,w)} = P_{vw} = \frac{1}{deg(v)}$ iff $(u, v), (v, w) \in E$ (once you get to node v, it doesn't matter how you got there), and $\forall (v, w) \in E$ we have

$$\sum_{u:((u,v),(v,w)) \in E'} P'_{(u,v)(v,w)} = \sum_{u:(u,v) \in E} \frac{1}{deg(v)} = 1$$

Applying the fact that G' is doubly stochastic means that Π' (the probability distribution across the nodes) is uniform to get that

$$\Pi'_{(v,u)} = \frac{1}{|V'|} = \frac{1}{2m}$$

so,

$$h'_{(v,u)(v,u)} = \frac{1}{\Pi'_{(v,u)}} = 2m$$

$\square$

Now, back to the main theorem.

Theorem 2. *For any undirected, connected, aperiodic* $G = (V, E)$*, we have* $\mathcal{C}(G) = O(mn) < O(n^3)$

Proof. Pick a starting vertex v_0 and build a spanning tree T of G rooted at v_0. Note that the number of edges in T is exactly $n-1$. Let $v_0, v_1, v_2, \ldots, v_{2n-2}$ be a depth first traversal of T. Notice that $v_{2n-2} = v_0$, and each edge of T appears exactly twice, once in each direction. So,

$$\begin{aligned}
\mathcal{C}(G) &\leq \sum_{j=0}^{2n-2} h_{v_j v_{j+1}} \\
&= \sum_{(u,v)\in T} C_{uv} \\
&\leq \sum_{(u,v)\in T} 2m \\
&= O(nm)
\end{aligned}$$

$\square$